

W3C

Beijing, China 2009-11-13

Klaus Birkenbihl, W3C klaus@w3.org

Introducing W3C

- W3C <http://www.w3.org/> the home of (X)HTML, XML, CSS, RDF, the Web and Semantic Web ...
- 400+ members (the usual suspects, SMEs, users, grassroots ...) <http://www.w3.org/Consortium/Member/List>
- ~65 groups doing the work <http://www.w3.org/Consortium/activities>
- 18 world offices all over the world <http://www.w3.org/Consortium/Offices/>
- a team of ~60 individuals - working and living distributed around the globe - coordinated by 3 hosts
 - MIT, US <http://www.csail.mit.edu/>
 - ERCIM, Europe <http://www.ercim.org/>
 - Keio University, Japan <http://www.keio.ac.jp/>
- we reach the community (4Mio hits/day on <http://www.w3.org/>)
- director: WWW inventor Tim Berners Lee <http://www.w3.org/People/Berners-Lee/>



W3C and Web business

- The Web is a multibillion \$s business
- There have to be rules based on consensus to have competitors work together
- Such rules are
 - *our process*<http://www.w3.org/Consortium/Process/>
 - *our liberal copyright*<http://www.w3.org/Consortium/Legal/2002/ipr-notice-20021231>
 - *our patent policy*<http://www.w3.org/Consortium/Patent-Policy/>
 - *CD, CI and what we say, e.g. logo usage*<http://www.w3.org/Consortium/Legal/logo-usage-20000308>
 - *neutrality, no product endorsement*
 - *or local point of contact is the W3C China Office*
<http://www.chinaw3c.org/>
- rules are liberal and leave space for creativity while maintaining the interest of our members and the public

The Web

After ~17 years in the wild the Web has changed and still develops in various directions:

- Web everywhere – the Mobile Web
- Web for everyone – the Accessible Web
- Web of structured information – the Semantic Web
- Web of participation and communication – Collections, Blogs and Fora
- Web as a service – Rich Web Applications
- Web for business – Web Services
- Web for media – Video and Voice on the web

Admitted: there might be different views on this

Requirements for the Web

Some general demands in order to make the Web useful and successful:

- privacy on the Web
- security and trust on the Web
- treatment of intellectual property
- content classification and rating
- accessibility of the Web
- scalability of the Web

It started pretty simple

- a common syntax for documents that may contain hyperlinks (HTML)
 - *provides a base for structuring documents*
 - *a framework for applications to display document*
 - *a mechanism to link to other documents*
- a syntax for locating documents (or fragments of documents) and the option to use this syntax within HTML to build hyperlinks (URL)
 - *a base for a globally unique identification of documents*
 - *unique identification it relies on the Internet Domain Name System*
- a protocol to send request and responses between clients (usually browsers) and servers (HTTP)
 - *to make servers and clients understand each other*

Fairly simple, but most of the topics mentioned before where not yet addressed - except: it scaled!

Broadening the Basis

- URL concept extended to URI and IRI
 - *a more general scheme for uniquely naming documents, real objects and beings, abstract concepts and ideas ...*
- HTML as basis is replaced by a more general concept: XML
 - *document structuring tags, can be freely chosen*
 - *can represent hierarchical databases*
 - *(apart from some syntactic restrictions of XML) HTML can be seen as a "restricted XML"*
 - *XHTML resolves this syntax problem*
- A few supporting technologies came with XML as there are:
 - *XML name spaces as a means to separate one space of tags from another*
 - *XML Schemas to allow the definition of restrictions to specify a new XML language*
 - *XQuery/Xpath as a means to navigate through or query XML documents*
 - *XML Base to cope with relative URIs/IRIs*
 - *XPointer to allow reference to XML documents and fragments*
 - *the DOM an abstract Interface to process XML and related events*

Even broader ...

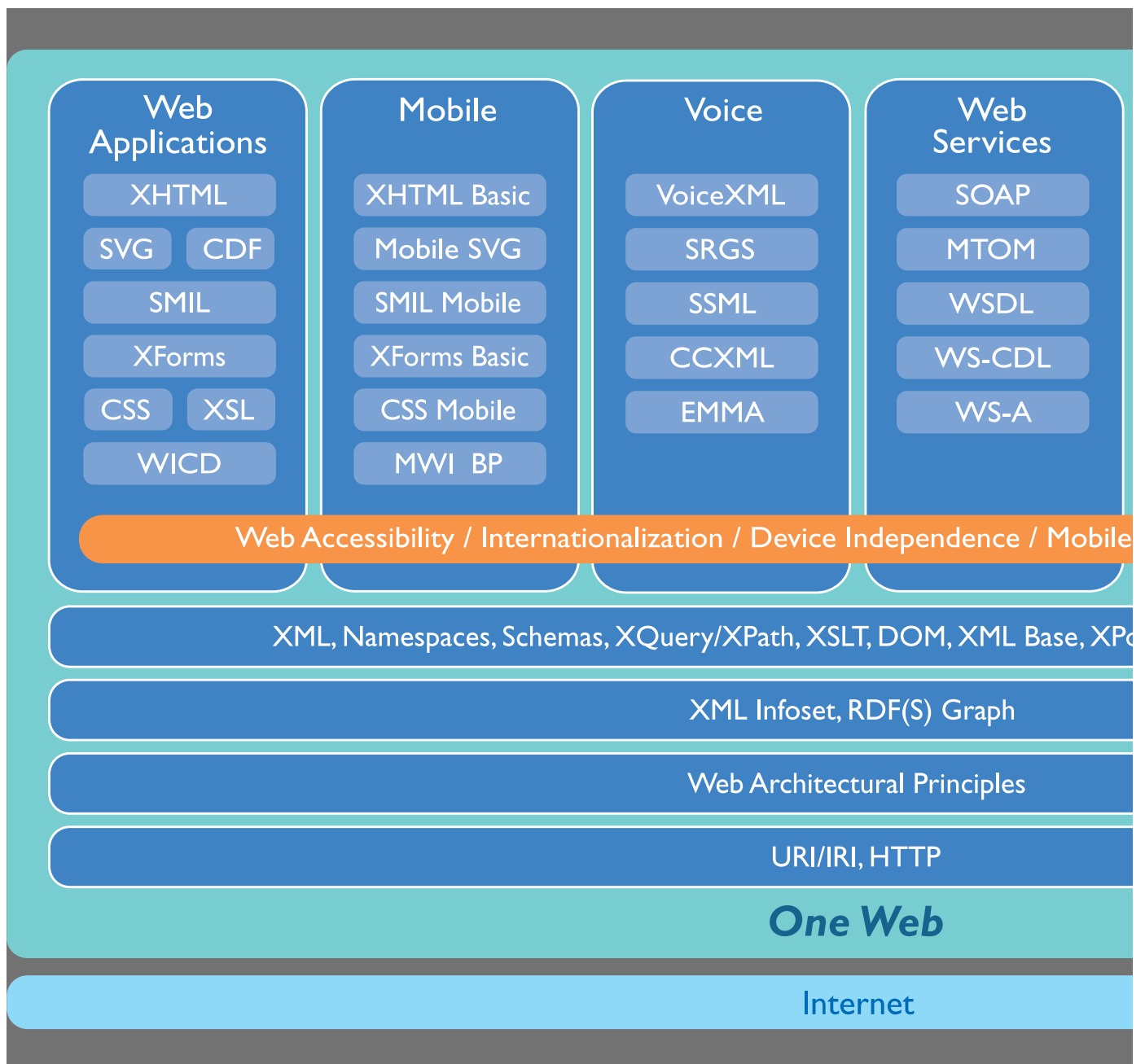
- XML relies on the concept to define complex entities by composing them of simpler ones.
- RDF graphs try to define entities by describing their relationship to others.

- *An RDF graph is composed of triples of the form subject predicate object*



- *subject (a resource) denotes something you want to define, predicate defines the relationship (property), and object (a resource) is something for which the relationship holds*
 - *Objects (and even predicates) in one triple can appear as subjects in others*
- As XML RDF comes with a set of associated technologies
 - *RDFS is a Schema language that defines a vocabulary that gives meaning to some resources and properties. This opens a standardized way to express set or logic related properties. OWL supplements and extends these definitions.*
 - *SPARQL is a query language for RDF*

Technology



We cannot cover all. More on <http://www.ict-media.de/webtech.html>

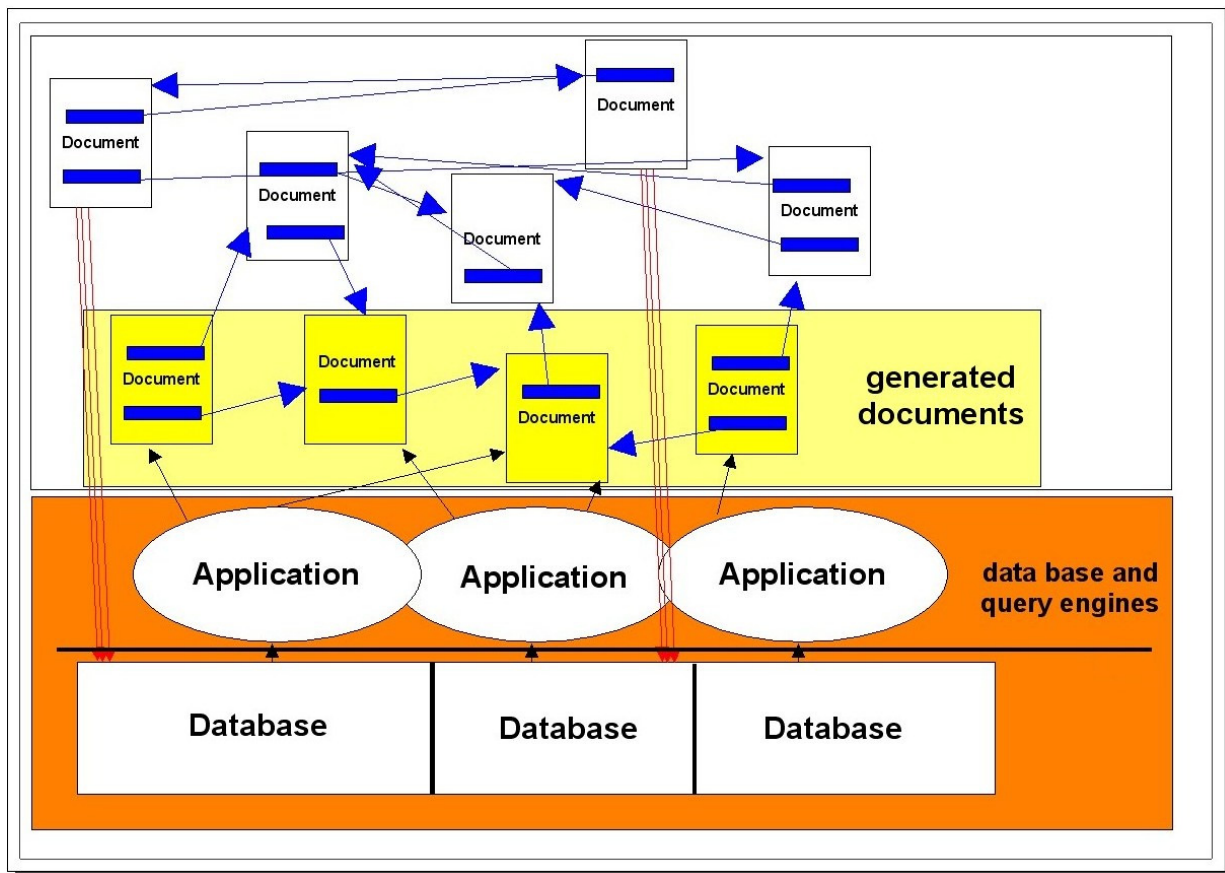
Web applications

- plain HTML is all you need to build a Web application
- on the other end we see fully fledged applications for:
 - *office work or picture processing.*
 - *Video, audio*
 - *sophisticated graphics*
 - *presentational software for mathematics or chemistry*
- Plugins and Interpreters for programming languages allow nearly any content or processing that works on the desktop. Using technologies that are
 - *standardized*
 - *proprietary*
 - *Open Source*
 - *locked*
- some important standards
 - *HTML ... 2 are emerging actually*
 - the main focus of XHTML 2 is to better use the advantages of XML within HTML
 - HTML 5 is more on advancing the existing HTML (HTML 4 or XHTML 1) in a compatible way
 - *CSS to control presentation in a structured way*
 - *Xforms as a better way to exchange information with servers*
 - *SVG for vector based two dimensional graphics*
 - *compound document standards (WICD, CDF...) to allow better integration of parts of a document using different technologies*
- The Web – especially when combined with powerful programming languages and interfaces – can act as a platform for

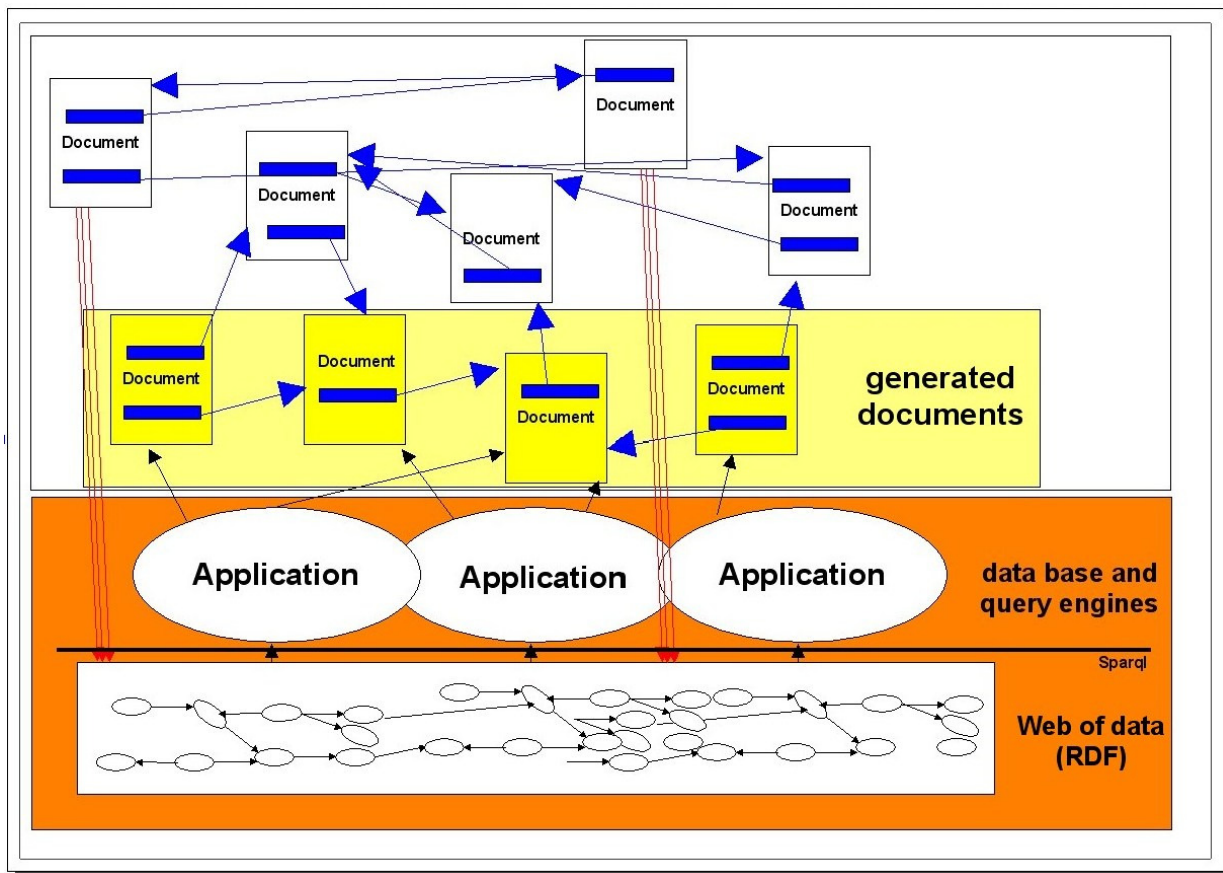
sophisticated applications.

- *Games, maps, image processing or complete office suites...*
 - *using technologies like Compound Document Formats, XML, the DOM (and DOM events), Xforms, XHTML, SVG and some more*
- SMIL to build presentations with media from various sources in a synchronized manner

The web today

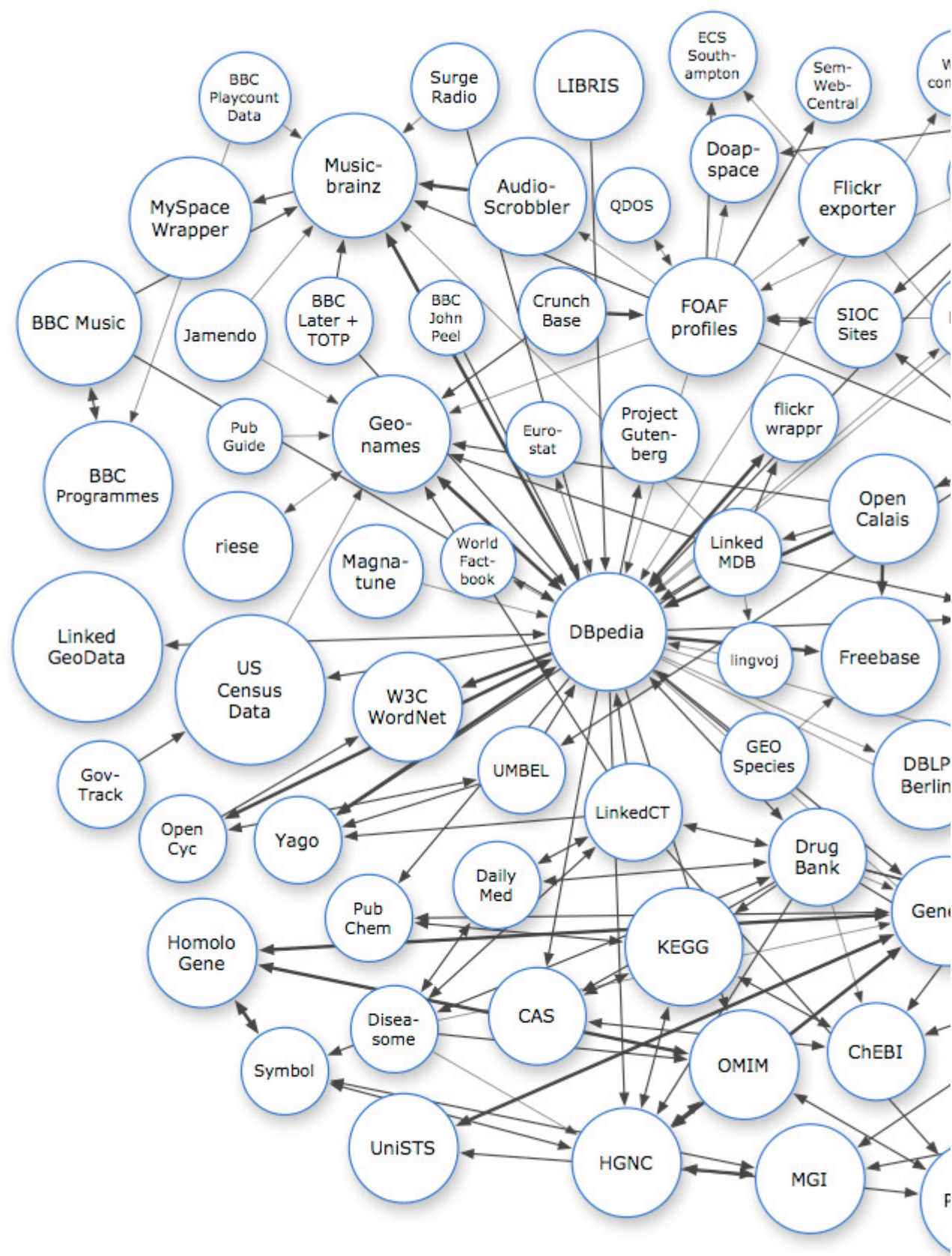


Semantic Web



How do we get there

- Stop: Mash-ups/Ajax do it already!
 - *API based, often one API per source*
 - *APIs mostly proprietary*
 - *quite some effort in implementing*
 - *dependency from API provider*
 - *instrumental to show how powerful a Web of data can be*
- Microformats:
 - *standard to represent data*
 - *could be used by mash-up*
 - *not very flexible, rather static*
- standards to resolve this:
 - *RDF family (RDFS, OWL) to represent data*
 - *SPARQL to access data*
 - *RDFa allows to include RDF statements in an XHTML document*
 - *SKOS provides a data model for sharing and linking knowledge organization systems as application of RDF*
- Linking Open Data gives an impressive example of how data can be linked



Mobile Web

- devices with limited capabilities (e.g. mobile phones)
 - *connectivity*
 - *bandwidth*
 - *processing power*
 - *display size and sound quality*
- limited capabilities are addressed by
 - *down sizing of existing standards to match the constraints of mobile devices*
 - *the development of guidelines especially for authors (but for authoring-tool manufacturers as well) on how to make sure that a Web site is mobile friendly*
 - *a way for servers to learn about the capabilities of an attached client*
- technologies:
 - *downsized standards XHTML Basic, SVG Tiny, SVG Basic, SMIL mobile, Xforms Basic and CSS mobile*
 - *DDR simple API describes an interface that servers can use to query device description*

Privacy and security

■ issues

- *while using the Web we disclose a lot of our personal data, our interests and preferences*
- *can we be sure that the Web site that pretends to be the home page of our bank is really our bank?*
- *can we securely transfer confidential data over the Web?*
- *can we securely identify ourselves on the web?*

■ technologie that help

- *P3P is a standard for specifying a machine readable description of a privacy policy*
- *TLS/SSL resolves much of the authentication, authorithation, confidentiality issues on the Internet transport layer*
- *XML Encryption, XML Signature and XML Key Management: 3 related technologies that allow to apply well known cryptographic methods on fragments of XML documents*

The Web live

■ here we are and ...

- *technologies and standards are a means to an end*
- *the list is far from complete*
- *available bandwidth and compute power are steadily increasing*
- *Web turned from geek's toy to a powerful infrastructure*
- *formed by millions of individuals, business people, journalists, artists, scientists...*
- *Google, Yahoo, Flickr, Ebay, Youtube, Myspace, Wikipedia, Mozilla are only a few popular projects that live in this one space*

■ here we go

- *a lot of open questions that wait for ideas, technical solutions, standards or a social agreement*
- *in parallel to the work on millions of Web sites, Blogs, Wikis, shops and fancy applications there is ongoing work on better standards and technologies*
- *W3C provides a framework where vendors, authors and users cooperate to improve the supporting frame of the Web*